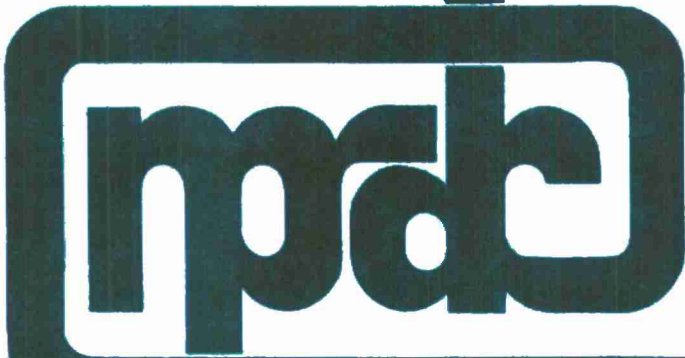


AD 780 515



NAVY PERSONNEL RESEARCH AND DEVELOPMENT CENTER, SAN DIEGO, CALIFORNIA 92152

*al*  
NPRDC *al* TR-74-28

FEBRUARY 1974

*"Illinois University.*

**RETENTION OF TEXT INFORMATION AS A FUNCTION  
OF THE NATURE, TIMING, AND NUMBER OF QUIZZES**

APPROVED FOR PUBLIC RELEASE,  
DISTRIBUTION UNLIMITED.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER TR 74-28	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) RETENTION OF TEXT INFORMATION AS A FUNCTION OF THE NATURE, TIMING, AND NUMBER OF QUIZZES		5. TYPE OF REPORT & PERIOD COVERED Technical Report 1 February - 1 September 73
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Richard C. Anderson      Peter M. Zych John R. Surber          Claire E. Lieberman W. Barry Biddle		8. CONTRACT OR GRANT NUMBER(s)  N61339-73-C-0078
9. PERFORMING ORGANIZATION NAME AND ADDRESS  University of Illinois Urbana, Illinois 61801		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS  Navy Personnel Research and Development Center San Diego, California 92152		12. REPORT DATE February 1974
		13. NUMBER OF PAGES 24
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report)  UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report)  Approved for Public Release; Distribution Unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES  Prepared in cooperation with Navy Personnel Research and Development Center.		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Prose learning Effects of questioning on recall Instructional technology		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) In two experiments a total of 662 high school students read a prose passage, took a verbatim or paraphrase quiz, and a week later completed a verbatim or paraphrase delayed test. Taking a quiz significantly enhanced performance on the delayed test. Performance was consistently much higher on the verbatim than the paraphrase forms of quizzes and tests. Fitting the data rather well was a theory which assumes that a verbatim question is best at evoking retrieval of phonologically coded information in short term memory whereas a		

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

paraphrase question is best at instigating transfer of the information into long-term, semantic memory.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

RETENTION OF TEXT INFORMATION AS A FUNCTION OF THE  
NATURE, TIMING, AND NUMBER OF QUIZZES

Richard C. Anderson  
John R. Surber  
W. Barry Biddle  
Peter M. Zych  
Claire E. Lieberman

University of Illinois at Urbana-Champaign

This research was supported by the  
Advanced Research Projects Agency  
of the Department of Defense and was  
monitored by Navy Personnel Research  
and Development Center under Contract  
No. N61339-73-C-0078.

The views and conclusions contained in this document are those of the  
authors and should not be interpreted as necessarily representing the  
official policies, either expressed or implied, of the Advanced Research  
Projects Agency or the U. S. Government.

Prepared for

Navy Personnel Research and Development Center  
San Diego, California 92152



## FOREWORD

This research was undertaken to examine problems of relevance to educational technology under a contract between NPRDC and the University of Illinois. It was an offshoot of the development of a Computer-Assisted Instruction Study Management System (CAISMS), and provides important information for instructors who wish to shape or mold the way in which their students process the textual material they're given to study. It is simple to recommend that instructors pose frequent questions to students about study material, but implementation of such a recommendation is not simple and requires information regarding the form and timing of the questioning. This research provides some of this information.



## Summary and Conclusions

### Problem

The purpose of this research was to study the direct effects of questioning on prose learning and retention. A primary issue, important to any instructional treatment, is whether answering a question about text causes a person to get the meaning from a communication or whether it causes him to learn the surface form of the message. A secondary issue was whether questions asked immediately or after a short delay have differential effects on remembering the material.

### Background

Numerous investigations have demonstrated the facilitative effects of questions during or shortly after instruction. However, since in every previous study verbatim form questions were used, it would be reasonable to suppose that the direct effect of questions entails little more than rote memorization of the form of the initial questions. If this is the case, the processing of the meaning of the text may be superficial. Several encoding stages are necessary to learn from text. First, perceptual features must be processed, then word strings are coded acoustically, and finally semantic processing must occur. Therefore, verbatim questions might involve minimal semantic processing and might be based primarily on phonological storage of the information. The use of questions constructed by paraphrasing should produce semantic processing of material. In addition, it is generally assumed that short-term memory involves primarily phonological, echoic storage of material and that this storage is transient, degrading with time. Long-term storage is presumably semantic in nature. Therefore, if time intervals intervene between reading text and questioning, answers should be based on any residual semantic storage rather than on the superficial orthographic and phonological characteristics of material.

### Approach

The experiments described in this paper compare the effects of verbatim and paraphrase questions on delayed retention of text information. The paraphrase questions presented after the passage were expected to occasion meaningful, semantic processing of the text information which resided, at the time, in short-term phonological storage. Verbatim questions presented after the passage would be processed in terms of their orthographic-acoustic features. That is to say, verbatim questions would demand less, if any, semantic encoding. Thus, students who received paraphrase questions were expected to perform better than those receiving verbatim questions. Since, short-term memory degrades with time delay, asking questions immediately or after interval was expected to degrade performance especially for verbatim materials.



In Experiment 1 240 subjects who were stratified into three levels of verbal ability, read one of two versions of a passage, completed a verbatim or paraphrase quiz, and a week later took a verbatim or paraphrase delayed test. Half the subjects took the quiz immediately after reading the passage, the remainder after 20 minutes.

In Experiment 2, run because of the results of the first experiment, the effect of taking a verbatim test and a paraphrase test on delayed retention was examined. Using essentially the same procedure and materials 422 subjects divided into groups read the passage, then completed a verbatim quiz, a paraphrase quiz, a verbatim quiz twice, a paraphrase quiz twice, a verbatim quiz followed by a paraphrase quiz or vice versa. One control neither read nor received the initial quiz.

### Results

Perhaps the most interesting and important finding was that on every occasion on which a quiz was given performance was better on the verbatim rather than the paraphrase form. The difference was greater if the quiz was given immediately after reading.

The studies were begun with the idea that a paraphrase quiz would lead to better delayed test performance than a verbatim quiz. This did not happen. What was overlooked in the original hypothesis was that phonological information in short-term memory is accessible to verbatim questions but relatively inaccessible to paraphrases.

### Recommendation

A revised theory was formulated which receives support from the data. Verbatim questions are more likely to allow retrieval of information from short-term memory whereas, once retrieved, a paraphrase question is likely to instigate transfer of the information into long-term, semantic memory. A verbatim quiz followed by a paraphrase quiz produced superior delayed retention. Also, consistent with the theory, was the fact that performance was higher on verbatim than paraphrase questions, especially when questions were answered immediately after reading the passage.

Therefore, it is recommended that the form and timing of questions used to guide and control the processing of the meaning in text material follow this prescription.

## CONTENTS

	Page
Foreword . . . . .	v
Summary and Conclusions . . . . .	vii
List of Tables . . . . .	x
Abstract . . . . .	xi
I. Introduction . . . . .	1
II. Experiment 1 . . . . .	4
A. Method . . . . .	4
1. Subjects . . . . .	4
2. Materials . . . . .	4
3. Design and Procedure . . . . .	4
B. Results . . . . .	5
C. Discussion . . . . .	6
III. Experiment 2 . . . . .	7
A. Method . . . . .	7
B. Results and Discussion . . . . .	8
IV. General Discussion . . . . .	9
References . . . . .	13
Distribution List . . . . .	15

## TABLES

Page

Mean Percentage Correct on the Verbatim and Paraphrase	
Forms of Quizzes and Tests . . . . .	5

RETENTION OF TEXT INFORMATION AS A FUNCTION OF THE  
NATURE, TIMING, AND NUMBER OF QUIZZES

Richard C. Anderson  
John R. Surber  
W. Barry Biddle  
Peter M. Zych  
Claire E. Lieberman

University of Illinois at Urbana-Champaign

Abstract

In two experiments a total of 662 high school students read a prose passage, took a verbatim or paraphrase quiz, and a week later completed a verbatim or paraphrase delayed test. Taking a quiz significantly enhanced performance on the delayed test. Performance was consistently much higher on the verbatim than the paraphrase forms of quizzes and tests. Fitting the data rather well was a theory which assumes that a verbatim question is best at evoking retrieval of phonologically coded information in short term memory whereas a paraphrase question is best at instigating transfer of the information into long term, semantic memory.

# RETENTION OF TEXT INFORMATION AS A FUNCTION OF THE NATURE, TIMING, AND NUMBER OF QUIZZES

## I. Introduction

That asking students questions will increase learning and retention is perhaps the best documented proposition there is in the field of instructional psychology. Studies reviewed by Gates (1917) indicate that the effects were already known just after the turn of the century. Gates' own research showed substantial benefits from "active recitation" on the learning and remembering of both serial lists of nonsense syllables and factual prose passages. Jones (1923) had subjects read three text selections and immediately thereafter complete one of two tests covering the selections. A day later all subjects took both tests. Scores on the repeated test were twice as high as scores on the test taken for the first time. Numerous investigators since then have corroborated the facilitative effects of questions during or shortly after instruction. Of course, the recent wave of interest in "test-like events" has been stimulated by the research of Ernst Rothkopf.

Both direct and indirect effects of questioning have been demonstrated. By a "direct effect" we mean the increment in performance which is observed when a question asked during or shortly after exposure to text is repeated later. The direct effect is usually large. It is not uncommon for the mean of the questioned group to be one and a half to two times the mean of a reading-only control group on repeated questions. Rothkopf (1966) can be credited with first showing that questions inserted within text after the sections to which they pertain also have indirect effects. On new post-test questions, unrelated to the inserted questions, groups that receive the inserted questions score higher than reading-only control groups. Since 1966 at least a half dozen experiments by several different investigators have confirmed that adjunct questions have small but consistent indirect effects (though see Ladas, 1973).

Even though the direct effects are invariably larger, it is the indirect effects of questioning which have tickled the imagination of research workers and captured the lion's share of their attention, perhaps precisely because the indirect effects are subtle and nonobvious. The purpose of this research was to study further the large and obvious direct effects of questioning.

The first issue is whether answering a question causes a person to get the meaning from a communication or whether, on the other hand, it merely causes him/her to learn the surface form of the message. This is an issue that ought to be raised with respect to any instructional treatment. It is especially relevant to an evaluation of the direct effects of questioning, since in every previous study, with which we are acquainted at any rate, the posttest has repeated the previous questions in literal, verbatim form. It would not be unreasonable to suppose that the direct effect of questions entails nothing more than rote learning of the orthographic-acoustic features of the initial questions. Half of the subjects

in the experiments reported herein received a posttest in which each question was a paraphrase of a question answered earlier. For these subjects, in other words, the posttest repeats the semantic content but not the lexical form of the initial questions. If the direct effect of questions is one of learning meanings these subjects will do better than control subjects on paraphrase as well as verbatim posttest questions, whereas to the extent that the direct effect of questions is simply a matter of learning surface forms the people who get the initial questions will have an advantage only on verbatim posttest questions.

A second issue investigated in the first experiment was the timing of initial questions. Either the entire set of questions, or quiz, was answered immediately after reading the passage or after a 20 minute filled delay. The working hypothesis was that a quiz question sets the occasion for mental review and further cognitive processing of text information. When the quiz question happens to make contact with information in short term memory, it is theorized that there is some probability that this information will be transferred into long term, semantic memory. Of course, a question would not be expected to affect information already in long term memory. Nor could a question influence information that had not been learned at all. The prediction was that people who received a quiz immediately would do better on the delayed test than people who completed the test after a 20 minute interval, because after an interval the information which potentially could have been affected by a quiz will have dropped out of short term memory.

We chose to depart from the now almost habitual practice of inserting a few questions after each of a series of small sections of text. Frase (1968) found little difference in performance on repeated posttest questions as a function of whether the questions had been answered initially after reading 10, 20, 40, or 50 lines of text (see also Boyd, 1973). In contrast, an earlier generation of educational psychologists found huge differences in delayed retention as a function of the timing of initial questions (Spitzer, 1939; Sones & Stroud, 1940); however, they studied intervals calibrated in hours, days, and even weeks instead of seconds and minutes. Hence, the critical interval might be larger than that investigated by Frase.

The third and final issue was the influence of the nature of the initial questions on delayed retention. With the caveat that many research reports don't provide enough information to be sure about the nature of the tests, it is still probably true that most demonstrations of the effects of questioning have tested recall of "facts" with items that repeated text statements in nearly verbatim form. Positive results have been obtained with such items, but there are grounds for arguing that items which required comprehension would give even stronger results. The argument goes like this. Several encoding stages are necessary to learn from text. First, the text is processed in terms of perceptual features. Second, the word strings are

coded according to acoustic features. Third, the meaning in the communication are brought to mind. These stages have been called orthographic, phonological, and semantic encoding, respectively (Anderson, 1972). There is general agreement that long-term memory is semantic in character (cf. Rumelhart, Lindsay, & Norman, 1972) though there is evidence that phonologically coded material can be remembered for a long time (see Posner, 1972). Short-term memory appears to entail phonological coding.

It is not inevitable that semantic encoding take place. A person may be under the impression he is "reading" when in fact all he is doing is saying the words to himself, without any contact with their potential meanings. There is now a really substantial case that learning is enhanced by procedures that cause people to semantically encode sentences rather than merely translate them into speech. For instance, Anderson and Kulhavy (1972) asked college students to study one sentence definitions of a series of unfamiliar words. Students who created and said aloud a sensible sentence containing each defined word did markedly better on a test of comprehension than students who read each definition aloud. Other evidence that procedures which induce meaningful processing facilitate learning from sentences and connected discourse has been reviewed in papers by Montague (1972) and Barclay (1973) as well as the one by Anderson and Kulhavy.

Several studies have obtained positive results with "higher order" questions (e.g., Berliner, et al., 1973); however, these studies have not included structurally identical questions that could be answered on the basis of surface features of the text. This criterion was met by Watts and Anderson (1971), who asked high school seniors to answer a question after reading each of five 450-word passages explaining a psychological principle. Subjects who received questions that required them to identify a new example of each principle performed significantly better overall on the posttest than all other subjects, including subjects who had answered otherwise identical questions that repeated examples described in the text. Especially noteworthy was the large advantage for new-example subjects on posttest questions which entailed still other new examples, different from any they had seen in the text or encountered in previous questions.

The experiments described in this paper compared the effects of verbatim and paraphrase questions on delayed retention of text information. The hypothesis was that paraphrase questions presented after the passage would occasion meaningful processing of the text information in short term, phonological storage. Verbatim questions presented after the passage could be processed again in terms of orthographic-acoustic features. In other words, verbatim questions would not demand semantic encoding. Therefore, it was expected that people who initially got paraphrase questions would perform better on the delayed test than people who initially received verbatim questions.



## II. Experiment 1

### A. Method

1. Subjects. Participating were 240 sophomores, juniors, and seniors from the high school in a farming community in central Illinois.

2. Materials. Two versions of a 550-word passage on the social behavior of the army ant were written. The versions were identical except for 15 important sentences or clauses which were judged to convey the main ideas of the text. Each important sentence (or clause) in one version was paraphrased in the other version; that is, it was written to be equivalent in meaning to the sentence in the first version but to contain no substantive words in common, except technical terms for which it was difficult to find synonyms. For instance, one version of the passage contained the following important sentence, To a great extent the colony's cohesion results from secretions from the queen that are attractive to the workers. The matched sentence in the other version was, The greatest factor in keeping the nest together is chemical odors from the queen that the workers find pleasant.

A multiple-choice test item was prepared for each important sentence. A segment of the sentence was removed. The remainder was transformed into a question. The deleted segment served as the correct answer alternative. To complete the item, three plausible wrong answer alternatives were invented. The items constructed for the matching sentences from the two versions of the passage were in one-to-one correspondence. Equivalent segments of the sentences served as correct response alternatives. The same three distractors were employed. The end result was that each form of the test contained verbatim items with respect to one of the versions of the passage and paraphrase items with respect to the other.

The Educational Testing Service Wide Range Vocabulary Test (French, Ekstrom, & Price, 1963) was used to measure verbal ability.

3. Design and procedure. The design entailed four orthogonal factors. Subjects stratified ex post facto into three levels of verbal ability read one of the two versions of the passage, completed a verbatim or paraphrase quiz, and a week later took a verbatim or paraphrase delayed test. Half of the subjects took the quiz immediately after reading the passage, the remainder after 20 minutes. During the delay subjects completed the Wide Range Vocabulary test and other verbal comprehension tests. These tests were completed following the quiz by subjects who received the immediate quiz.

The experiment was run in the school cafeteria in two shifts of about 120 subjects. Subjects were assigned to conditions simply by distributing randomly-ordered stacks of booklets containing the experimental materials.



Instructions mimeographed on the first page of the booklet stated that the passage should be read carefully, that no notes should be taken, that the reader should stop at the end of the passage or when told to stop, and that a test would be given. Instructions preceding the quiz emphasized that no one should look back at the passage. Color coding of the pages in the experimental booklet made it easy for the three assistants who were monitoring the experiment to make sure this direction was followed. Not even the teachers were informed that a delayed test would be given a week later. Ample time was allowed for every subject to complete every phase of the experiment.

## B. Results

Overall, students who received a quiz averaged 61.6% on the delayed test whereas those who did not receive one averaged 54.5%,  $t(157) = 3.16$ ,  $p < .01$ . However, two specific hypotheses could not be confirmed in the form outlined earlier. First, students who took the immediate quiz averaged 63.5% on the delayed test as compared to 59.7% for the people who completed the quiz 20 minutes after reading the passage, which is not a significant difference,  $F(1,125) = 2.34$ ,  $p = .129$ . Second, to our considerable surprise, the delayed test mean for the group that got a verbatim quiz was actually higher than the mean for the group that received a paraphrase quiz, though not significantly so,  $F(1,125) = 2.26$ ,  $p = .135$ .

Performance was substantially better on the verbatim than the paraphrase quiz,  $F(1,125) = 22.87$ ,  $p < .01$ , and also better on the verbatim than the paraphrase delayed test,  $F(1,181) = 5.63$ ,  $p < .05$ . As can be seen from Table 1, the difference between verbatim and paraphrase forms diminished somewhat over time.

TABLE 1

Mean Percentage Correct on the Verbatim and Paraphrase  
Forms of Quizzes and Tests

Occasion	Form	
	Verbatim	Paraphrase
Immediate Quiz	80.4%	63.7%
20 Minute Quiz	71.8	63.2
Delayed Test (No Quiz)	58.5	50.6
Delayed Test (Quiz)	63.7	59.5

There was a significant Form of Quiz X Form of Test interaction,  $F(1,125) = 6.40, p < .05$ . Scores on the delayed test were highest when the test form matched the quiz form. Also appearing, however, in the analysis of the quiz scores, was a significant Time of Quiz X Form of Quiz X Form of Test interaction,  $F(1,125) = 5.79, p < .05$ . Since the delayed test had not been administered yet, this must mean that the groups were not initially equivalent, presumably because of a random perturbation. When differences among groups on the quiz were discounted, the interaction involving delayed test scores almost disappeared.

Verbal ability affected both quiz performance,  $F(2,125) = 30.34, p < .01$ , and delayed test performance,  $F(2,181) = 26.86, p < .01$ . Also obtained were significant effects for Passage and Passage X Ability. Since these effects are not germane to the main issue confronted in this paper they will not be discussed.

### C. Discussion

We had confidently expected the paraphrase quiz to facilitate delayed test performance more than the verbatim quiz. The fact that the trend of the results ran in the opposite direction caused us to revise our theory. It is now argued that the process by which a quiz question enhances delayed retention involves two stages. First, the question must permit retrieval of information from short term memory. Second, the question must instigate meaningful processing of the information so as to transfer it into long term, semantic memory. This theory can be expressed in the following equation,

$$P(D) = k + (1-k)rt,$$

which says that the probability of a correct response on the delayed test,  $P(D)$ , equals the proportion of items of information the person already has in long term storage,  $k$ -- that is, information he knows before reading or learns from reading the passage, plus an increment due to taking the quiz. The increment consists of that information not already known which the questions cause the person to retrieve,  $r$ , from short term, phonological memory and transfer,  $t$ , into long term, semantic memory.

Below are estimators for the three parameters, letting  $C$  be the mean proportion correct of the no-quiz control group, and  $Q_j$  and  $F_j$  be the proportion correct for the  $j$ th group (or person) on the quiz and delayed (final) test, respectively.

$$\hat{k} = C$$

$$\hat{r}_j = \frac{Q_j - C}{Q_j - C}$$

$$\hat{t}_j = \frac{F_j - C}{Q_j - C}$$

The parameters were calculated for the present data using simply the mean quiz and delayed test scores, pooling over the two forms of the test. Obtained were values of  $r$  of .48 and .20 and values of  $t$  of .42 and .58 for the verbatim and paraphrase quiz groups, respectively. From the perspective of the model there is support for our original contention that paraphrase questions would be better than verbatim questions at promoting the transfer of information into long term storage. On the other hand, verbatim items proved vastly better at evoking retrieval of the information to begin with. It is not difficult to understand why. Short term memory is largely phonological in character. A matching phonological string can be produced given a verbatim cue but not a paraphrased one.

The augmented theory further illuminates the effects to be expected from the timing of questions. Only verbatim questions can tap a large proportion of the information in short term memory; so, the length of the interval between reading the passage and answering the question should be important only for performance on verbatim questions. In the analysis already reported which failed to show an effect due to the timing of the quiz, the verbatim and paraphrase quiz groups had been pooled. The picture changes when just the groups that received a verbatim quiz are considered. There was a significant difference between immediate and 20 minute verbatim quiz scores,  $t(85) = 2.57$ ,  $p < .05$ . Also significant was the difference in delayed test scores of the groups that had received immediate and 20 minute verbatim quizzes,  $t(85) = 2.33$ ,  $p < .05$ .

### III. Experiment 2

A main purpose of the second experiment was to test a nonobvious prediction from the theory developed during the post mortem on the first experiment: an optimum treatment should be a verbatim quiz followed by a paraphrase quiz. According to the augmented theory a verbatim question allows retrieval of phonologically coded information in short term memory. Thus primed, the information is more likely to be accessible for meaningful processing instigated by the paraphrase question. The net result should be an increased likelihood that the information will get into permanent storage.

#### A. Method

Four hundred and twenty-two freshmen from a suburban high school read the army ant passage and then completed a verbatim quiz, a paraphrase quiz, a verbatim quiz twice, a paraphrase quiz twice, a verbatim quiz followed by a paraphrase quiz, or paraphrase quiz followed by a verbatim quiz. One control group neither read the passage nor took a quiz. Another control group read the passage but did not receive a quiz. In all other respects the design, materials, and procedure were the same as in Experiment 1.

## B. Results and Discussion

The initial analysis of the data involved four planned, orthogonal comparisons. It was expected, first, that reading the text would improve delayed test performance. The grand mean for groups exposed to the text was 52.9%. The mean for the groups that took the test without an opportunity to study the passage was 34.0%. This difference was significant,  $t(398) = 8.28$ ,  $p < .01$ .

Second, as expected, students who received a quiz averaged higher on the delayed test than students who did not receive a quiz,  $t(398) = 1.71$ ,  $p < .05$ , though the difference was not very great, just 53.6% for the quiz groups as compared to 48.9% for the no-quiz group.

Third, the groups that completed two quizzes had a delayed test mean of 55.8% while the mean for the groups that completed one quiz was 49.2%, an advantage for the two-quiz condition which had been predicted,  $t(208) = 2.53$ ,  $p < .01$ . However, it happens that the groups that received one quiz performed worse (though not significantly so) on the first quiz than did the two-quiz groups. This means the groups were not equivalent to begin with. Thus, it remains to be seen whether two quizzes are actually better than one.

The fourth and most interesting prediction was that students who received a verbatim quiz followed by a paraphrase quiz would do better on the delayed test than students who received other combinations of two quizzes. The prediction was confirmed,  $t(138) = 1.73$ ,  $p < .05$ . The verbatim-paraphrase group averaged 60.2% on the delayed test whereas the other two-quiz groups averaged 54.3%.

One element of the argument that a verbatim quiz followed by a paraphrase quiz would maximally enhance delayed test performance was that the verbatim quiz would "prime" information so as to make it more accessible to the paraphrase questions. This contention was tested directly. The performance of the verbatim-paraphrase group on the paraphrase quiz was compared to the performance of the three groups that began with a paraphrase quiz. The latter groups averaged 56.7% whereas the verbatim-paraphrase group averaged 61.6%. While in the predicted direction, the difference fell short of being significant,  $t(208) = 1.41$ .

The parameters of the model outlined earlier were computed for the groups that received just one kind of quiz. (It was not clear how to proceed with the mixed quiz conditions.) The values of  $r$  and  $t$  were .34 and .24, respectively, for the verbatim and verbatim-verbatim groups, and .16 and .31 for the paraphrase and paraphrase-paraphrase groups. These figures show the same trends as the ones obtained in the first experiment.

Students averaged 68.1% on the first verbatim quiz but only 56.7% on the first paraphrase quiz,  $F(1,208) = 19.01$ ,  $p < .01$ . Similarly, the mean on the verbatim form of the delayed test was 57.2% whereas the mean on the paraphrase form was 50.0%,  $F(1,208) = 7.75$ ,  $p < .01$ . Performance was only

slightly higher on the verbatim than the paraphrase second quiz, however,  $F(1,138) = 1.44$ . As already indicated, scores on the paraphrase quiz improved when preceded by a verbatim quiz. On the other hand, a paraphrase first quiz actually led to lower scores on a subsequent verbatim quiz.

As in the first experiment, students did somewhat better on the delayed test when the form matched the form of the quiz, though the interaction was not significant,  $F(1,208) = 3.04$ ,  $p = .083$ .

Verbal ability affected performance on the first quiz,  $F(1,208) = 24.54$ ,  $p < .01$ , the second quiz,  $F(1,208) = 28.17$ ,  $p < .01$ , and the delayed test,  $F(1,208) = 19.09$ ,  $p < .01$ . The only other  $F$  significant at the .05 level was for a five-way interaction involving the second quiz.

#### IV. General Discussion

Perhaps the most interesting and important finding of the present research was that on every occasion on which a quiz or test was given performance was better on the verbatim than the paraphrase form. The difference was greatest on the quiz given immediately after reading. Smaller but still significant differences appeared in both experiments on the test given a week after reading.

Discounting guessing and previous knowledge, the assumption is that a person can answer a paraphrase question only if he has semantically encoded the relevant text information, whereas a verbatim item can be answered if the information has been encoded either semantically or phonologically (see Anderson, 1972). The second experiment included a control group that answered the questions without reading the passage. The difference of the scores of this group and the scores on the immediate quiz of the groups that did read the passage gives the amount of information acquired from reading, which was 34.2% for the group that got the immediate verbatim quiz but just two thirds as large, 22.7%, for the group that got the immediate paraphrase quiz. The conclusion is that one third of the "knowledge" that resulted from reading the passage depended upon asking questions which reinstated the exact language of the text. In simple, old fashioned terms there was evidently a lot of rote learning going on.

Is there any escape from this conclusion? One alternative interpretation can be ruled out on the basis of the design and procedures. The fact that performance was higher on verbatim than paraphrase forms of quizzes and tests definitely cannot be attributed to the differential difficulty of the two forms, since the versions of the passage and forms of the quizzes and tests were counterbalanced. What was a paraphrase item for one subject was a verbatim item for the next.



A second possibility is that any text information that could be recalled at all was semantically coded, but that many of the paraphrases were inadequate considering the linguistic naivety of the subjects. A roughly paraphrased question might fail to serve as a cue for text information even if the information were semantically coded. This interpretation cannot be definitely ruled out on the basis of data in hand. There are counter indications, however. First, in both experiments scores declined more sharply from immediate quiz to delayed test on the verbatim than on the paraphrase form. This fact is consistent with the dual coding theory, which says that the memorial half life of phonologically coded information is less than that of semantically coded material, but inconsistent with the notion that all the information was semantically coded. Second, if the roughness of paraphrase interpretation were correct, an interaction between verbal ability and form of quiz or test would have appeared. Students with high verbal ability would have done about as well on either verbatim or paraphrase quizzes or tests because they would have been sophisticated enough to see the semantic equivalence of the two forms. By the same line of reasoning, there would have been a marked difference between verbatim and paraphrase scores for students of low ability. The fact is that there was no suggestion of such an interaction in the present experiments.

Research is in progress to check the equivalence of the two forms of the questions. In the meantime we judge the dual coding hypothesis provides a better explanation of the results than does an argument based on the assumption that the paraphrase was inadequate.

Without known exception, the identical questions have been repeated in the previous studies showing a direct questioning effect. As was pointed out in the introduction, this fact leaves open the possibility that the effect is trivially specific. The present research was designed to see if the effect would appear when the items in the final test repeated the semantic content but not the lexical form of the initial questions. Delayed test scores were significantly higher in Experiment 1 when the test form matched the quiz form. However, the groups were not equivalent to begin with and the interaction largely disappeared when initial differences were discounted. There was a nonsignificant trend in Experiment 2 for delayed test scores to be higher when the test form matched the quiz form.

The failure to find a dependable interaction does support the conclusion that the direct effect of questions involves more than the learning of the surface form of the quiz questions, but only weakly so. The support would have been stronger if the groups that received verbatim quizzes were to have done better than the control group on the paraphrase delayed test and the groups that received paraphrase quizzes were to have done better than the control group on the verbatim delayed test. This did not happen. Students who completed different forms of the quiz and test were not significantly better than the control subjects in either experiment. Therefore, to sum up, all we are able to say is that the present experiments do not permit a firm conclusion as to whether the direct effect of questions depends upon repeating the literal wording of the questions.

We began with the idea that a paraphrase quiz would lead to better delayed test performance than a verbatim quiz. This hypothesis turned out badly. What we had overlooked in formulating the hypothesis was that information in short term memory is phonologically coded, making it accessible to a verbatim question but relatively inaccessible to a paraphrased one.

The augmented theory turned out well. The data from both experiments supported the view that a verbatim question is more likely to allow retrieval of information from short term memory whereas, once retrieved, a paraphrase question is more likely to instigate transfer of the information into long term, semantic memory. The second experiment confirmed a prediction from the theory: a verbatim quiz followed by a paraphrase quiz optimumly facilitated delayed retention. Consistent with the theory, finally, was the fact that performance was higher on verbatim than paraphrase questions, especially when the questions were answered immediately after reading the passage.

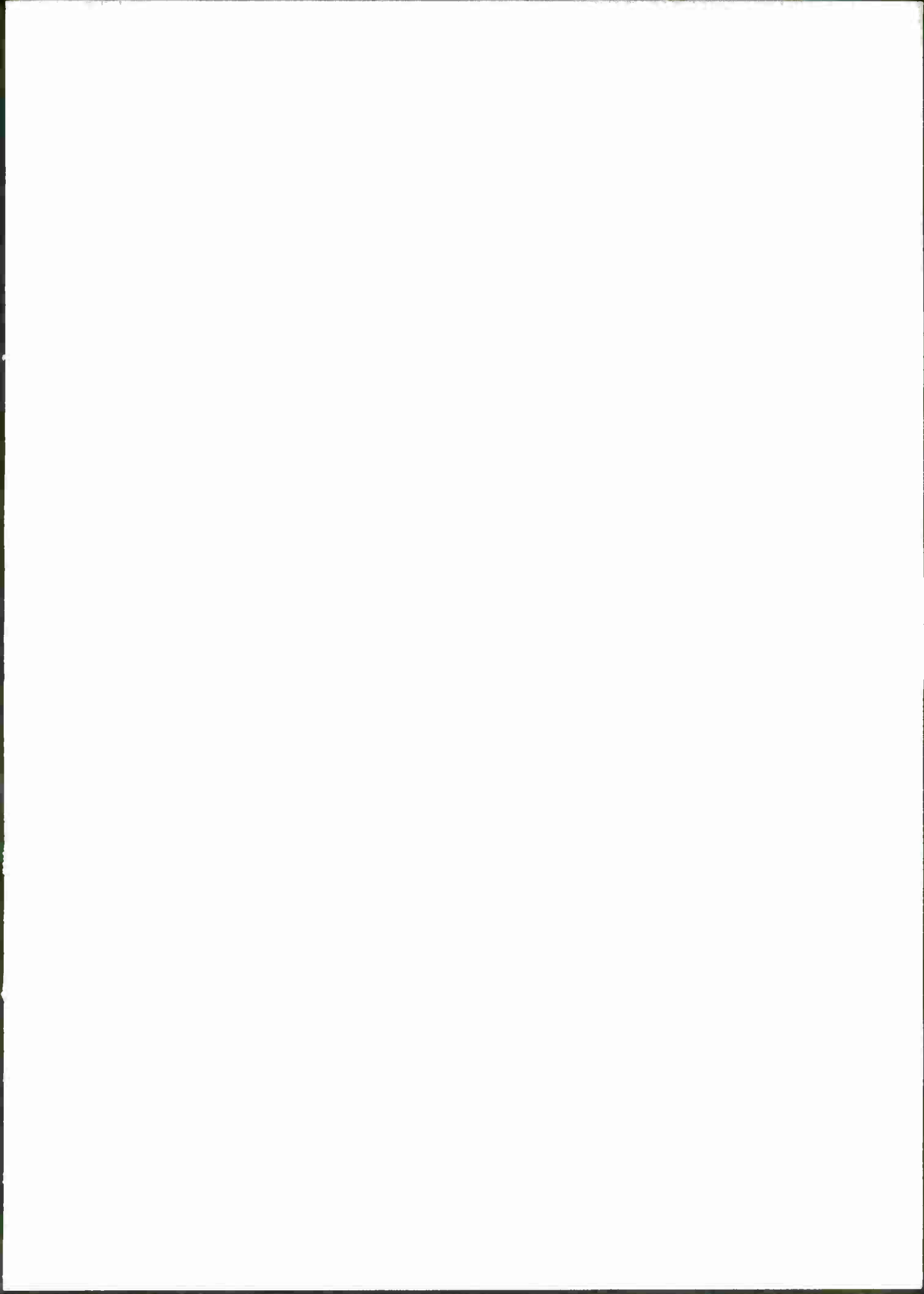




## References

- Anderson, R. C. How to construct achievement tests to assess comprehension. Review of Educational Research, 1972, 42, 145-170.
- Anderson, R. C. & Kulhavy, R. W. Learning concepts from definitions. American Educational Research Journal, 1972, 9, 385-390.
- Barclay, J. R. The role of comprehension in the remembering of sentences. Cognitive Psychology, 1973, 4, 229-254.
- Berliner, D. C., Shavelson, R. J., Ravitch, M. M., & Loeding, D. Individual differences in the effects of adjunct questions on learning from prose material. Paper presented at the meetings of the American Educational Research Association, New Orleans, February 25 - March 2, 1973.
- Frase, L. T. Effects of question location, pacing, and mode of retention of prose material. Journal of Educational Psychology, 1968, 59, 244-249.
- French, J. W., Ekstrom, R. B., & Price, L. A. Kit of reference tests for cognitive factors. Princeton, N. J.: Educational Testing Service, 1963.
- Gates, A. I. Recitation as a factor in memorizing. Archives of Psychology, 1917, 26, (Whole No. 40).
- Jones, H. E. The effect of examination on permanence of learning. Archives of Psychology, 1923-24, 10, 1-70.
- Ladas, H. The mathemagenic effects of factual review questions on the learning of incidental information: A critical review. Review of Educational Research, 1973, 43, 71-82.
- Montague, W. E. Elaborative strategies in verbal learning and memory. In G. H. Bower (Ed.), The psychology of learning and motivation, Vol. 6. New York: Academic Press, 1972.
- Posner, M. I. & Warren, R. E. Traces, concepts, and conscious constructions. In A. W. Melton and E. Martin (Eds.), Coding processes in human memory. Washington, D. C.: V. H. Winston & Sons, 1972.
- Rothkopf, E. Z. Learning from written instructive materials: An exploration of the control of inspection behavior by test-like events. American Educational Research Journal, 1966, 3, 241-249.
- Rumelhart, D. E., Lindsay, P. H., & Norman, D. A. A process model for long-term memory. In E. Tulving and W. Donaldson (Eds.), Organization of memory. New York: Academic Press, 1972. Pp. 198-245.

- Sones, A. M. & Stroud, J. B. Review with special reference to temporal position. Journal of Educational Psychology, 1940, 31, 665-676.
- Spitzer, H. F. Studies in retention. Journal of Educational Psychology, 1939, 30, 641-656.
- Watts, G. H. & Anderson, R. C. Effects of three types of inserted questions on learning from prose. Journal of Educational Psychology, 1971, 62, 387-394.



## DISTRIBUTION LIST

Chief of Naval Operations (OP-987E)  
Chief of Naval Personnel (Pers-10c)  
Chief of Naval Education and Training (CNET N-33)  
Chief of Naval Technical Training  
Chief of Naval Material (NMAT-030B)  
Chief of Naval Research (Code 458) (2)  
                                (Code 450) (4)  
Commanding Officer, Naval Training Equipment Center  
Commanding Officer, Navy Medical Neuropsychiatric Research Unit  
Commanding Officer, Naval Aerospace Medical Institute  
Superintendent, Naval Academy  
Superintendent, Naval Postgraduate School  
Superintendent, United States Military Academy  
Superintendent, Air Force Academy  
Superintendent, Coast Guard Academy  
Army Research Institute for Behavioral and Social Sciences  
Personnel Research Division, Air Force Human Resources Laboratory (AFSC),  
    Lackland Air Force Base  
Occupational Research Division, Air Force Human Resources Laboratory (AFSC),  
    Lackland Air Force Base  
Technical Training Division, Air Force Human Resources Laboratory (AFSC),  
    Lowry Air Force Base  
National Research Council  
National Science Foundation  
Science and Technology Division, Library of Congress  
Defense Documentation Center (12)



U16074